

Pekrun, Reinhard

## **Vergleichende Evaluationsstudien zu Schülerleistungen: Konsequenzen für die Bildungsforschung**

*Zeitschrift für Pädagogik 48 (2002) 1, S. 111-128*



Quellenangabe/ Reference:

Pekrun, Reinhard: Vergleichende Evaluationsstudien zu Schülerleistungen: Konsequenzen für die Bildungsforschung - In: Zeitschrift für Pädagogik 48 (2002) 1, S. 111-128 - URN: urn:nbn:de:0111-opus-38234 - DOI: 10.25656/01:3823

<https://nbn-resolving.org/urn:nbn:de:0111-opus-38234>

<https://doi.org/10.25656/01:3823>

in Kooperation mit / in cooperation with:

# **BELTZ**

<http://www.beltz.de>

### **Nutzungsbedingungen**

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

### **Terms of use**

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

### **Kontakt / Contact:**

peDOCS  
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipt.de](mailto:pedocs@dipt.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

## Inhaltsverzeichnis

### *Thema: Grundlagenforschung in der LehrerInnenbildung*

*Barbara Koch-Priewe*

Grundlagenforschung in der LehrerInnenbildung.

Einführung in den Thementeil..... 1

*Georg Hans Neuweg*

Lehrerhandeln und Lehrerbildung im Lichte des Konzepts des

impliziten Wissens..... 10

*Matthias Schierz/Jörg Thiele*

Hermeneutische Kompetenz durch Fallarbeit. Überlegungen

zum Stellenwert kasuistischer Forschung und Lehre an Beispielen

antinomischen Handelns in sportpädagogischen Berufsfeldern..... 30

*Karin Nölle*

Probleme der Form und des Erwerbs unterrichtsrelevanten

pädagogischen Wissens..... 48

### *Diskussion: Leistungsvergleiche im Schulwesen*

*Dietrich Benner*

Die Struktur der Allgemeinbildung im Kerncurriculum moderner

Bildungssysteme. Ein Vorschlag zur bildungstheoretischen Rahmung

von PISA..... 68

*Ewald Terhart*

Wie können die Ergebnisse von vergleichenden Leistungsstudien

systematisch zur Qualitätsverbesserung in Schulen genutzt werden?.... 91

*Reinhard Pekrun*

Vergleichende Evaluationsstudien zu Schülerleistungen: Konsequenzen für zukünftige Bildungsforschung .....	111
---	-----

*Besprechungen*

*Klaus Prange*

Alfred K. Trembl: Allgemeine Pädagogik. Grundlagen, Handlungsfelder und Perspektiven der Erziehung Jürgen Oelkers: Einführung in die Theorie der Erziehung .....	129
--	-----

*Fritz Osterwalder*

Diane Ravitch: Left Back. A Century of Failed School Reform .....	138
---	-----

*Daniel Tröhler*

Bijan Adl-Amini: Pestalozzis Welt. Eine Einladung zur Erziehung .....	142
---	-----

*Christian Lüders*

Siegfried Müller/Heinz Sünker/Thomas Olk/Karin Böllert (Hrsg.): Soziale Arbeit. Gesellschaftliche Bedingungen und professionelle Perspektiven .....	145
---	-----

*Dokumentation*

Pädagogische Neuerscheinungen .....	148
-------------------------------------	-----

## Content

### *Topic: Basic Research in Teacher Education*

*Barbara Koch-Priewe*

Basic Research in Teacher Education – An introduction ..... 1

*Georg Hans Neuweg*

Teacher Action and Teacher Education in the Light of the Concept  
of Tacit Knowing..... 10

*Matthias Schierz/Jörg Thiele*

Hermeneutic Competence Through Case Studies. Considerations  
on the relevance of casuistic research and theory exemplified by  
antinomic acting in the field of sports pedagogics ..... 30

*Karin Nölle*

Problems of both Form and Aquisition of Pedagogical Knowledge  
Relevant for Teaching ..... 48

### *Discussion: Large Scale Assessments*

*Dietrich Benner*

The structure of General Education in the Core Curriculum  
of Modern Educational Systems. A proposal for a framework for  
PISA based on educational theory ..... 68

*Ewald Terhart*

How Can the Results of Large Scale Assessments Be Used Systematically  
for an Improvement of the Quality of Schools? ..... 91

*Reinhard Pekrun*

Large Scale Assessments: Implications for further research ..... 111

Book Reviews ..... 129

New Books..... 148

Reinhard Pekrun

## Vergleichende Evaluationsstudien zu Schülerleistungen: Konsequenzen für die Bildungsforschung

**Zusammenfassung:** Internationale und nationale Vergleichsstudien zu schulischen Leistungen liefern Daten zu Verteilungen, Strukturen und Korrelaten von Kompetenzen und Leistungen bei Schülern. Sie dienen damit einem deskriptiven Systemmonitoring des schulischen Bildungswesens. Ungeachtet mancher Erwartungen von Politik und Öffentlichkeit aber lassen sie aufgrund ihrer meist querschnittlichen Untersuchungsdesigns kaum Schlüsse auf die Bedingungen und Wirkungen von Schülerleistungen oder auf pädagogische Handlungsmöglichkeiten zu. Diskutiert wird, welche Desiderata für zukünftige Bildungsforschung aus den Möglichkeiten und Grenzen solcher Studien zu folgern sind. Abzuleiten ist, dass auch in Deutschland in verstärktem Maße entwicklungs-, bedingungs- und interventionsanalytisch orientierte Forschungsprogramme zur schulischen Bildung implementiert werden sollten. Ferner ergeben sich Folgerungen für die Methodik und Organisation solcher Forschungsprogramme.

### 1. Ausgangslage

#### 1.1 Welchen Ertrag liefern Vergleichsstudien zu Schülerleistungen?

Aktuelle Vergleichsstudien zu Schülerleistungen wie die *Third International Mathematics and Science Study* (TIMSS) und das *Program for International Student Assessment* (PISA) zeigen in konsistenter Weise, dass Schülerleistungen in beträchtlichem Ausmaß variieren. Dies gilt für Leistungsunterschiede zwischen Nationen ebenso wie für Unterschiede innerhalb von Nationen. Die intranationale Variation ist in einigen Fällen kaum geringer als die internationale Variation (ein Beispiel sind die erheblichen Leistungsunterschiede zwischen Schulen in den USA). Das durchschnittliche Abschneiden deutscher Schüler ist in Fächern wie Mathematik und Naturwissenschaften eher dem Mittelfeld wirtschaftlich entwickelter Nationen zuzuordnen, wobei signifikante Variationen zwischen Bundesländern, zwischen Schulen und Schulklassen innerhalb von Bundesländern und zwischen Einzelschülern in Rechnung zu stellen sind (vgl. Baumert/Bos/Lehmann 2000; Baumert/Lehmann 1997; Weinert 2001).

Über die Ermittlung des Durchschnitts und der Variation von Leistungen in den untersuchten Schülerpopulationen hinaus liefern solche Studien Evidenz zu den internen Strukturen der jeweiligen Leistungsbereiche (also z.B. zu den Strukturen mathematischer Leistungen) und zu Verknüpfungen zwi-

schen diesen Bereichen (z.B. zwischen mathematischen und sprachlichen Leistungen). Ferner bestätigen die Befunde, dass Schülerleistungen mit anderen Variablen korrelativ verknüpft sind, für die angenommen werden kann, dass sie auf Leistungen Einfluss nehmen oder von ihnen abhängen. Bestätigungen ergeben sich insbesondere für Individualvariablen des Schülers (z.B. Intelligenz, Motivation) und für proximale Umweltvariablen (vor allem Merkmale des schulischen Fachunterrichts).

Die Variablenkonfiguration der Studien ist zunächst jeweils am vorliegenden Forschungsstand zu Schülerleistungen und ihren Bedingungen orientiert. Im Sinne aktueller Desiderata zu den Bildungswirkungen von Schule wird in Teilbereichen aber auch Neuland deskriptiv erschlossen. Dies gilt heute insbesondere – z.B. bei PISA – für die Erhebung von Ausprägungen und Strukturen fächerübergreifender ‚Schlüsselqualifikationen‘ (*„cross-curricular competencies“*; vgl. Klieme u.a. 2001; Weinert 1998). Beispiele sind Fähigkeiten zum Problemlösen, Kompetenzen zum selbstregulierten Lernen und soziale Kompetenzen von Schülern. Auch hier handelt es sich dann jeweils um Befunde zu Verteilungen, internen Strukturen und externen Korrelaten entsprechender Kompetenz- und Leistungswerte.

Studien wie TIMSS, PISA etc. liefern damit ein deskriptives Systemmonitoring des Bildungswesens. Sie ermitteln Charakteristika der Verteilungen und Strukturen von Schülerleistungen und sind insoweit – aus der Perspektive einer Deskription des Bildungswesens – outputorientiert. Darüber hinaus erfassen sie in deskriptiver Weise den Stand von Input-, Prozess- und Kontextvariablen schulischer Lern- und Leistungsprozesse sowie korrelative Verknüpfungen zwischen Schülerleistungen und solchen Variablen.

## 1.2 Was können solche Vergleichsstudien nicht leisten?

Die öffentliche Diskussion der Resultate von TIMSS und ähnlichen Studien zeigt, dass an solche Studien häufig Erwartungen gerichtet werden, die über die genannten Deskriptionsleistungen hinausgehen.

Mit der Durchführung solcher Studien wird in der Regel unterstellt, dass die erfassten Lernleistungen prädiktiv sind für individuelle und kollektive Leistungen jenseits des schulischen Bereichs (insbesondere in nachschulischer Ausbildung und Beruf): Leistungen von Schülern interessieren, weil sie für die zukünftige wirtschaftliche und kulturelle Leistungsfähigkeit von Nationen und Regionen als bedeutsam angesehen werden. Ferner wird erwartet, dass aus Verknüpfungen zwischen Leistungen und anderen Variablen auf Bedingungen des Zustandekommens von Leistungen geschlossen werden kann. So wird häufig angenommen, dass Unterschiede des Fachunterrichts in unter-

schiedlichen Nationen ursächlich für Unterschiede von Schülerleistungen im jeweiligen Fach seien (vgl. Weinert 2001). Schließlich wird darüber hinaus häufig auch erwartet, dass aus solchen Verknüpfungen nicht nur auf Leistungsursachen, sondern auch auf *Handlungsmöglichkeiten* zur Leistungsoptimierung und Prävention von Defiziten geschlossen werden könne.

Tatsächlich aber können Erwartungen dieser Art von Vergleichsstudien wie TIMSS und PISA kaum erfüllt werden. Angesichts der Untersuchungsanlage solcher Studien sind Aussagen zu prognostischen Beziehungen, Bedingungsbeziehungen und Handlungsmöglichkeiten aus den Daten dieser Studien selber nicht ableitbar: Bei Vergleichsstudien zu Schülerleistungen handelt es sich in der Regel um summative Evaluationen mit querschnittlichem, nicht-experimentellem Design, also einer Designart mit geringer kausaler Aussagekraft.

Stellt man beispielsweise fest, dass ein bestimmtes didaktisches Vorgehen im Mathematikunterricht (z.B. in japanischen Schulen) mit höheren Schülerleistungen verknüpft ist als andere Vorgehensweisen (z.B. in US-amerikanischen und deutschen Schulen), so ist eine solche korrelative Verknüpfung unterschiedlichen Interpretationen gegenüber offen: (a) Unterschiedliche Unterrichtsformen könnten unterschiedliche Schülerleistungen bedingen; (b) unterschiedliche Leistungsniveaus von Schülern könnten Lehrern unterschiedlich anspruchsvolles Unterrichten ermöglichen; (c) sowohl Unterrichts- wie Leistungsunterschiede könnten auf Unterschiede der Schülermotivation zurückgehen, ohne untereinander verknüpft zu sein; etc.

Dies ist methodisch trivial, wird aber auch in der Fachdiskussion häufig nicht hinreichend beachtet. Erst unter Zuhilfenahme von Bedingungswissen über Unterrichtseffekte und das Zustandekommen von Lernleistungen würde es möglich, eine korrelative Verknüpfung dieser Art kausal zu interpretieren. Schließen lässt sich zwar, dass eine Bedingungs-dynamik am Wirken war, welche die betreffende Verknüpfung produziert hat, nicht aber, um welche Dynamik es sich gehandelt hat. Ferner lassen sich aus dem Gefüge korrelativer Verknüpfungen Hypothesen zu möglichen Bedingungsbeziehungen ableiten und auf ihre Plausibilität hin testen (z.B. anhand synchron-bedingungsanalytischer Strukturgleichungsmodellierungen; vgl. Kline 1998). Ohne Zusatzannahmen aber sind Schlüsse auf Ursache-Wirkungs-Beziehungen auch unter Verwendung solcher Methoden nicht möglich. Dementsprechend lassen sich auch Handlungsempfehlungen ohne Zusatzwissen nicht ableiten, denn solche Empfehlungen setzen valide Einsichten in die Wirkungen von Handlungen voraus.

Hieraus folgt auch, dass gefundene Leistungsunterschiede zwischen Nationen, Bundesländern etc. nicht ohne Zusatzannahmen als Resultat unterschiedlicher Leistungsfähigkeit der jeweiligen Bildungssysteme interpretiert

werden können; vielmehr könnte es sich z.B. auch um differenzielle Leistungsfähigkeit von Elternhäusern handeln. Ausschließen ließe sich ohne Zusatzwissen noch nicht einmal die mancherorts diskutierte Hypothese, dass solche Leistungsunterschiede auf Unterschiede in der genotypischen Ausstattung von Mitgliedern unterschiedlicher Nationen zurückgehen könnten (vgl. Lynn 1997).

Für eine Einschätzung der bildungspolitischen Konsequenzen differenzieller Leistungsdaten dürften Möglichkeiten zu Rückschlüssen auf Wirkungen von Unterricht und Schulorganisation und auf Möglichkeiten der Optimierung von zentraler Bedeutung sein. Das notwendige kausale Zusatzwissen steht bisher kaum hinreichend zur Verfügung; folglich sind solche Rückschlüsse gegenwärtig nur eingeschränkt möglich.

## **2. Konsequenzen für zukünftige Forschungsprogramme**

Da querschnittlich-deskriptive Leistungsstudien nur bedingt zur Aufklärung von Leistungsfolgen, Leistungsbedingungen und Handlungsmöglichkeiten taugen, bedürfen solche Studien der gezielten Vertiefung durch bedingungs- und interventionsanalytisch angelegte Forschungsprogramme. Ziel solcher Programme sollte es sein, zu einer nachhaltigen Steigerung und Sicherung der Qualität unseres Bildungswesens beizutragen. Angesichts der Komplexität des Gegenstands ist dieses Ziel anhand von Einzelstudien nicht zu erreichen. Vielmehr bedarf es einer so hinreichend breit und systematisch angelegten Entwicklung der zuständigen Forschungsdisziplinen, dass eine Gewinnung kumulativer, untereinander vernetzter Erkenntnisse möglich wird.

Betrachtet man historische Entwicklung und gegenwärtigen Entwicklungsstand der beteiligten Disziplinen, wird deutlich, dass dies bisher nur in wenigen Teilbereichen der Fall war. In der Lehr-Lern-Forschung finden sich Bemühungen dieser Art seit den 70er-Jahren u.a. in BLK-Programmen und Schwerpunktprogrammen der DFG (wie z.B. dem seit April 2000 laufenden DFG-Programm „Die Bildungsqualität von Schule im mathematisch-naturwissenschaftlichen Unterricht“). Vergleicht man solche Programme in Struktur und finanzieller Ausstattung mit Forschungsprogrammen zu anderen für die Zukunft unserer Gesellschaft relevanten Problemen, wird deutlich, dass eine Intensivierung der Bildungsforschung in Deutschland erforderlich ist. Dies gilt insbesondere auch angesichts der Offensiven zur Bildungsforschung, die gegenwärtig in anderen Ländern zu verzeichnen sind (wie z.B. dem Programm *ROLE* der National Science Foundation in den USA).



Welche Formen und Ansatzpunkte von Forschung sind notwendig, wenn eine Qualitätssteigerung und -sicherung des Bildungswesens erreicht werden soll? Programme der Bildungsforschung lassen sich u.a. differenzieren nach (a) Gegenstandsbereichen (Institutionen, Fächer, Entwicklungsbereiche etc.); (b) Art der angezielten Aussagen (Beschreibung, Erklärung, Prognose, Interventionsgestaltung, Evaluation); (c) verwendeten wissenschaftlichen Methoden (Survey- vs. Einzelfallforschung, quantitative vs. qualitative Forschung etc.); und (d) zuständigen Fachdisziplinen (Fachdidaktiken, Pädagogik, Psychologie etc.). Eine differenzierte Taxonomie von Forschungsprogrammen soll hier nicht erstellt werden. Vielmehr soll die Diskussion auf diejenigen Bereiche konzentriert werden, für die aus meiner Sicht vordringlicher Forschungsbedarf besteht. Eingegangen wird zunächst auf inhaltliche Desiderata, anschließend auf Fragen der Methoden und Organisation zukünftiger Bildungsforschung.

### 3. Desiderata zu den Inhalten schulbezogener Bildungsforschung

#### 3.1 Strukturen und Entwicklungsverläufe des Wissenserwerbs

Die Lern-, Gedächtnis-, Wissens- und Expertisepsychologie hat erfolgreich untersucht, in welcher Weise Menschen Wissen erwerben. Die folgenden vier für die Wirkungen schulischen Lernens zentralen, über Basisprozesse des Wissenserwerbs hinausgehenden Fragen aber sind bisher erst ansatzweise geklärt (vgl. Gruber/Ziegler 1996; Klix/Spada 1998; Reusser 1998).

- 1) *Strukturen von Wissen.* Wie sind schulisch erwerbbarer Wissensbestände strukturiert? Ungeklärt ist diese Frage insbesondere für ‚Schlüsselqualifikationen‘ in Bereichen wie Problemlösen und soziale Kompetenzen. So ist beispielweise offen, zu welchen Anteilen es sich bei solchen Qualifikationen jeweils um domänenübergreifende Fähigkeiten (im Sinne von ‚cross-curricular competencies‘) oder aber um bereichsgebundene, domänenspezifische Fähigkeiten handelt, deren Erwerb in bereichsgebundener Weise erfolgen müsste (vgl. auch Weinert 1998).
- 2) *Entwicklungsverläufe des Wissenserwerbs.* In welcher kumulativen oder nicht-kumulativen Weise entwickeln sich Wissensbestände, die schulisch erworben werden? Diese Frage ist auch für fachgebundenes Lernen nicht hinreichend geklärt (Beispiel: Entwicklung von mentalen Grundvorstellungen in Mathematik und Naturwissenschaften jenseits des Grundschulalters), noch weniger allerdings für den schulischen Erwerb von fachübergreifenden Kompetenzen (z.B. Ceci 1991).

- 3) *Verhältnis von fachlichem zu fachübergreifendem Lernen.* In welchem Verhältnis stehen fachgebundene und fachübergreifende Lernprozesse? Auf diese Frage können wir bisher keine zureichenden Antworten geben. Dies gilt auch für die spezifische Frage, zu welchen Anteilen fachliches Lernen durch fachspezifische und fachübergreifende Prinzipien des Lernens zu erklären ist. Zu den Folgen zählt, dass das Verhältnis zwischen domänenübergreifenden Disziplinen wie Psychologie und Pädagogik einerseits und den domänenspezifisch orientierten Fachdidaktiken andererseits wissenschaftssystematisch bisher kaum hinreichend bestimmbar ist.
- 4) *Transfereigenschaften von Wissen.* Unter welchen Bedingungen zeigt schulisch erworbenes Wissen Transfer in außerschulischen Anwendungssituationen, und unter welchen Bedingungen bleibt es ‚träge‘? Bis heute wird diese Frage in der Situationismusdebatte sowohl grundsätzlich wie auch bezogen auf schulisches Lernen kontrovers diskutiert (z.B. Anderson u.a. 1997; Cobb/Bowers 1999; Gerstenmaier/Mandl 1995). Antworten wären von zentraler Bedeutung für Einschätzungen, welche Alltags- und Berufsrelevanz unterschiedlichen Arten von schulischem Lernen zukommt und welche Formen von Lehr-Lern-Prozessen im schulischen Bildungswesen realisiert werden sollten.

Forschung zu diesen vier miteinander verknüpften Problembereichen ist bisher in einer Reihe von teils voneinander isolierten Forschungstraditionen betrieben worden, zu denen u.a. Intelligenzpsychologie, Problemlösepsychologie, Expertiseforschung und empirisch-pädagogische Untersuchungen zu Schlüsselqualifikationen zählen. Zukünftige Forschung sollte interdisziplinär organisiert sein, um wissenspsychologischen und gegenstandsspezifischen Gesichtspunkten gleichermaßen Rechnung tragen zu können. Zentral müsste dabei eine Kooperation von Psychologie bzw. empirischer Pädagogik einerseits und empirisch-fachdidaktischer Forschung andererseits sein.

### *3.2 Bedingungen von Wissenserwerb und Schülerleistungen*

Welche Unterrichts-, Individual- und Kontextbedingungen sind für Wissenserwerb und Leistungen von Schülern verantwortlich? Antworten auf diese Frage sind nicht nur aus der Perspektive pädagogischer Grundlagenforschung wesentlich, sondern auch unter Handlungsperspektiven, da Optimierungen von Unterrichts- und Umweltgestaltung und schülerbezogene Interventionen geeignetes Bedingungswissen voraussetzen. Zu einigen Bedingungsbereichen liegt heute kumulative, konsistente Evidenz vor. Dies gilt für eine Reihe von Individualbedingungen (insbesondere kognitive Variablen wie Intelligenz

und Vorwissen) und für einige Unterrichtsvariablen (z.B. Lehr-Lern-Zeiten, Variablen direkter Instruktion). In anderen Bereichen hingegen sind zentrale Fragen bisher nicht hinreichend geklärt.

- 1) *Schulischer Unterricht*. Relativ gut erforscht sind die Bedeutung von Lehr-Lern-Zeiten, die Wirksamkeit bestimmter Variablen direkter Instruktion sowie einige Aspekte kooperativen Lernens. Hingegen ist nicht zureichend geklärt, welche Wirkungen Unterrichtsformen jenseits direkter Instruktion entfalten und in welcher Weise die Effekte unterschiedlicher Unterrichtsformen untereinander und mit Schülermerkmalen interagieren (vgl. Helmke/Weinert 1997).
- 2) *Lernen mit neuen Medien*. Von Wirtschaft und Öffentlichkeit werden Forderungen erhoben, auch im schulischen Bereich verstärkt auf Möglichkeiten des Lernens mit neuen Medien zurückzugreifen. Entscheidungen über ein zielgerichtetes Vorgehen in diesem Bereich bedürfen einer wissenschaftlichen Fundierung. Wir benötigen systematische Forschung zu den Wirkungen und Nebenwirkungen schulischen Lernens mit neuen Medien (Mandl u.a. 1998). Dabei ist auch in diesem Forschungsbereich eine Intensivierung der interdisziplinären Kooperation zwischen grundlagenorientierten Disziplinen und den jeweiligen Fachdidaktiken überfällig.
- 3) *Schülerbedingungen: Emotion, Motivation, Interesse*. Im Bereich der Schülerbedingungen des Wissenserwerbs hat sich die Forschung auf kognitive Faktoren konzentriert. Aus theoretischer Perspektive ist anzunehmen, dass affektiv-motivationale Variablen (Interesse, Lernfreude, Anstrengungsbereitschaft etc.) nicht weniger wesentlich sind. Die Forschung zu solchen Variablen ist disparat, sie bedarf der Verstärkung und interdisziplinären Zusammenarbeit. Anknüpfen lässt sich an die in Deutschland relativ starken Traditionen der Leistungsmotivations-, Interessen- und Emotionsforschung (vgl. Krapp 1993; Pekrun/Hofmann 1999; Schiefele, im Druck).
- 4) *Schülerkontexte: Elternhaus*. Die vorliegende Forschung zeigt, dass dem Elternhaus über die gesamte Schulzeit hinweg eine überragende Entwicklungsbedeutung zukommt, und zwar auch in direkt schulbezogenen Lern- und Entwicklungsbereichen (Pekrun, im Druck). Enge Beziehungen mit schulischem Wissenserwerb zeigen sich insbesondere für lern-, leistungs- und schulbezogene Verhaltensweisen der Eltern, hingegen weniger für distale Variablen des sozialen Kontextes (Helmke/Weinert 1997). Offen ist aber, inwieweit es sich bei solchen Beziehungen um Epiphanomene genetischer Kovariation von Eltern und Kindern handelt, oder ob tatsächlich Einflüsse des Elternverhaltens zugrunde liegen. Antworten auf diese Frage wären grundlegend für eine Einschätzung und Ge-

staltung elternbezogener Interventionen und eine gezielte Verbesserung der Kooperation von Elternhaus und Schule. Notwendig wäre hierfür eine bisher fehlende, gleichzeitig populationsgenetisch orientierte und schulbezogene Entwicklungsforschung, die ähnliche Überraschungen erbringen könnte wie die jüngere populationsgenetische Forschung zu familiären Einflüssen auf die Persönlichkeitsentwicklung (z.B. Loehlin u.a.1998).

- 5) *Schülerkontexte: Peergruppe und interkulturelle Probleme.* Probleme der Kooperation von Angehörigen verschiedener Kulturen und einer kulturellen Integration werden heute u.a. in der Arbeits- und Organisationspsychologie empirisch erforscht (vgl. Thomas 1996). Die schulbezogene Bildungsforschung sollte sich ebenfalls verstärkt um dieses Problemfeld kümmern.
- 6) *Schul-, Organisations- und Systementwicklung.* Fragen einer eigenständigen Schul- und Organisationsentwicklung durch Lehrer, Schulen und regionale Schulverwaltungen werden heute in Gestalt vieler lokaler Initiativen teils innovativ behandelt (vgl. Rolff 1996). Hingegen mangelt es an systematischer Forschung und Evaluation. Hilfreich wären in diesem Bereich Kooperationen mit der Organisationspsychologie und -soziologie.

### 3.3 *Folgen des Wissenserwerbs: Prognose- und Wirkungsforschung*

Für eine Einschätzung der Wirksamkeit von Schule sind Langzeituntersuchungen erforderlich, die den Einfluss von schulischen Institutionen und schulisch vermitteltem Wissenserwerb auf nachschulische Wissens- und Persönlichkeitsentwicklungen und zugeordnete Berufs- und Lebenserfolge analysieren. Da sich für Schulzeit und schulischen Unterricht einerseits und Erwachsenenalter sowie Arbeitswelt andererseits in Pädagogik, Psychologie und Soziologie jeweils unterschiedliche Teildisziplinen zuständig fühlen, mangelt es an Studien dieser Art.

In Erweiterung der oben genannten psychologischen Fragestellungen wäre in solchen Untersuchungen zu klären, in welcher Weise schulischer Wissenserwerb zu nachschulischem Lernen beiträgt und zu Wissenstransfer im nachschulischen Berufs- und Alltagsleben führt. Im Vergleich zu den Methodenmöglichkeiten einer auf die Schulzeit beschränkten Forschung ist bei der Realisierung entsprechender Forschungsprogramme ein erhöhter Aufwand in Rechnung zu stellen, da Individualstichproben (anstelle von Klassen- oder Schultichproben) zu ziehen wären, die zudem über lange Zeiträume begleitet werden müssten (mindestens bis zum Zeitpunkt eines feststellbaren Erwerbs beruflicher Expertise).

### *3.4 Interventionsforschung: Entwicklung, Implementierung und Evaluation von Bildungsmaßnahmen*

Für die Weiterentwicklung des Bildungswesens reicht eine Analyse der vorfindlichen Bedingungen und Wirkungen von Lern- und Leistungsentwicklungen nicht aus. Hierfür gibt es zwei wesentliche Gründe: (a) Es dürfte eher unwahrscheinlich sein, dass die Bandbreite wirksamer Bildungsmaßnahmen durch die historisch jeweils realisierten Formen von Maßnahmen bereits ausgeschöpft ist. Vielmehr ist anzunehmen, dass manche Möglichkeiten zu wirksamen Bildungsmaßnahmen erst durch gezielte Konstruktion gefunden werden. Dies gilt gegenwärtig z.B. für Formen des Lernens mit neuen Medien. (b) Der Versuch einer Implementierung von Bildungsmaßnahmen kann auf Probleme stoßen, zu denen eine lokale Begrenztheit ihrer Wirksamkeit ebenso zählen kann wie Widerstände seitens der Akteure. Unterricht, der an bestimmten Orten hoch wirksam ist, kann seine Wirksamkeit anderenorts verlieren, und Realisierungen sind auf Kooperation der Beteiligten angewiesen (Schulleitungen, Lehrer, Eltern etc.). So könnte es beispielsweise sein, dass eine Übertragung von Formen des japanischen Mathematikunterrichts auf deutsche Schulen an den spezifischen Motivations- und Disziplinarkonstellationen deutscher Schulklassen scheitert.

Möglichkeiten einer wissenschaftsbasierten Gestaltung und Evaluation bildungsbezogener Maßnahmen werden in Deutschland noch nicht hinreichend genutzt. Auch hierfür ist das bereits zitierte Lernen mit neuen Medien ein gutes Beispiel. In diesem Bereich dominiert heute vielerorts praxeologisches Vorgehen (z.B. in der Entwicklung von Lernsoftware) ohne differenzierte Begleitforschung. Soweit evaluiert wird, geht dies über eine Nutzung informeller, formativ-qualitativer Verfahren kaum hinaus (vgl. Tergan 1998; Wottawa/Thierau 1998). Dementsprechend stoßen Möglichkeiten mediengestützten Lernens in der Schule mancherorts auf Euphorie und anderenorts auf ablehnende Skepsis, ohne dass eine hinreichende Erkenntnisbasis für Schlussfolgerungen zu Gestaltungsformen, Anlässen, Lernerbedingungen, Situationsbedingungen und Nebenwirkungen vorhanden ist. Zielführende Entscheidungen über flächendeckende Realisierungen im schulischen Kontext lassen sich so noch nicht treffen.

Notwendig ist es also, systematische Interventionsforschung zu betreiben. Bestandteile solcher Forschung sollten Gestaltung, Implementierung und Evaluation von bildungsbezogenen Maßnahmen sein. Neben Unterrichts- und Schulinterventionen verdienen Programme zur Verbesserung der Kooperation von Schule und Elternhaus und zu einer angemessenen Nutzung von Evaluationsresultaten besonderes Augenmerk.

- 1) *Unterricht und Schulorganisation.* Aus normativen Überlegungen zu Bildungszielen einerseits und vorliegenden empirischen Erkenntnissen zu Bedingungen und Wirkungen bereits realisierter Bildungsmaßnahmen andererseits sind schrittweise Maßnahmen (*‘treatments‘*, *‘Programme‘*, *‘Interventionen‘*) zur Optimierung von Unterricht, schulischer Erziehung und Schulorganisation zu entwickeln. Vor einer flächendeckenden Einführung in einzelnen Regionen oder Ländern sollte jeweils in systematischer Weise als bisher Interventionsforschung anhand von Modellversuchen durchgeführt werden, die Möglichkeiten der Implementierung erproben und Wirkungen formativ und summativ evaluieren. Bei einer solchen Evaluation wäre nach den vorliegenden Erfahrungen neben einer Erfolgs- und Nebenwirkungskontrolle vor allem auch eine systematische, empirische Kontrolle der tatsächlichen Implementierung des betreffenden Programms in den beteiligten Schulen erforderlich (z.B. Strittmatter/Bedersdorfer 1991).

Interventionsforschung dieser Art kann auf unterschiedlichen Systemebenen angesiedelt sein und unterschiedliche Generalisierungsansprüche verfolgen. Es kann sich um Forschung mit universellen Generalisierbarkeitszielen handeln, aber auch um Forschung, deren Anspruch zunächst auf die lokalen Spezifika einer einzelnen Region oder einzelnen Schule zugeschnitten und damit begrenzt ist (z.B. schulspezifische Gestaltung und Evaluation eines Unterrichtsprogramms an einer einzelnen Schule). Universell und lokal orientierte Interventionsforschung stehen im Spannungsverhältnis eines Bandbreite-Fidelitäts-Dilemmas. Beide Formen der Forschung können sich aber fruchtbar ergänzen: Universelle Forschung liefert allgemeine Prinzipien, deren Erklärungs- und Wirkungsertrag durch lokale Forschung angereichert werden kann (vgl. auch Schibeci/Grundy 1987).

- 2) *Kooperation von Schule und Elternhaus.* Interaktionen zwischen Schule und Elternhaus sind in Deutschland häufig durch quantitative und qualitative Defizite gekennzeichnet. Familien aber nehmen entscheidenden Einfluss auf alle Kernbereiche von Entwicklung; dies schließt schulbezogene Entwicklungen ein (also u.a. den schulisch vermittelten Wissenserwerb), und zwar nicht nur in der Kindheit, sondern auch im Jugendalter. Verbesserungen der Kooperation zwischen beiden Sozialisationsinstanzen könnten sich als entscheidender Motor einer Optimierung schulischer Bildungswirkungen erweisen. Bisher mangelt es in Deutschland an entsprechenden Interventionsprogrammen und an zugeordneter Interventionsforschung (vgl. Pekrun 1997).
- 3) *Nutzung von Evaluationsergebnissen.* Entscheidend für eine Umsetzung von Evaluationsergebnissen in Gestalt einer Modifikation und Implemen-

tierung evaluierter Programme sind Rückmeldungen an die jeweiligen Akteure. Wir benötigen Forschung zu der Frage, wie Rückmeldungen zu gestalten sind und durch sie angestoßene Prozesse so zu begleiten sind, dass Programmmodifikation bzw. -implementierung in ergebnisadäquater Form erfolgen (Terhart 2002).

### *3.5 Forschung zu Lehrerbildung und schulbezogener Personalentwicklung*

Für eine Qualitätssicherung des schulischen Bildungswesens sind Lehrer, Schulleiter und Schulverwaltungen als entscheidende Eingangsgrößen anzusehen. Auch unter Perspektiven der Fortentwicklung von Schule bedürfen Lehrerbildung und schulbezogene Personalentwicklung in Deutschland der Reform: Veränderungen und Optimierungen schulischer Bildungsmaßnahmen setzen Akteure voraus, deren Berufshandeln durch Innovationsbereitschaft und wissenschaftsorientierte berufliche Fähigkeiten bestimmt ist (Terhart 2000). Als Grundlage von Reformen werden Evaluationen des gegenwärtigen Standes von Lehrerbildung und Personalentwicklung benötigt, und Reformmaßnahmen sollten ihrerseits ebenfalls empirisch evaluiert werden. Unsere bildungswissenschaftliche und -politische Praxis wird dem bisher noch nicht hinreichend gerecht. Mit TIMSS, PISA etc. wurden Maßnahmen einer systematischen Evaluation im primären und sekundären Bildungsbereich in die Wege geleitet. Entsprechendes sollte in den Bereichen der tertiären und berufsbegleitenden Lehrerbildung und der Personalentwicklung von Schulverwaltungen geschehen. Wir bedürfen in Deutschland einer Intensivierung der tertiären und berufsbezogenen Bildungsforschung; dies schließt Forschung zur Lehrerbildung ein (vgl. Perry/Smart 1997).

### *3.6 Erweiterung der Perspektiven I: Bildung, Persönlichkeitsentwicklung und Gesundheit*

Unter einer pädagogischen Perspektive dienen Lernprozesse nicht nur dem Erwerb von Wissen, das in Beruf und Alltag direkt verwertbar ist, sondern sollten sich darüber hinaus auch an allgemeineren Zielen der schulischen Bildung orientieren. Inwieweit schulische Bildungsmaßnahmen zur Realisierung solcher Ziele beitragen, lässt sich aus den Erträgen bisheriger Leistungsstudien nur begrenzt erschließen. In Anknüpfung an die ältere Forschung zu den Bildungs- und Persönlichkeitswirkungen von Schule bedarf es m.E. sowohl einer neuen normativen Bildungsdiskussion wie auch einer intensiveren Einbeziehung von Entwicklungsbereichen jenseits des Wissenserwerbs (Ent-

wicklung von Werten, politischen Einstellungen, Identität etc.; vgl. Benner 2002; Fend/Stöckli 1997; Pekrun/Fend 1991). Ferner schiene es mir wichtig, neben den Wissens- und Bildungswirkungen von Schule mögliche Nebenwirkungen zu berücksichtigen, also beispielsweise die gesundheitlichen Folgen schulischer Bildungsmaßnahmen.

Unter der normativen Perspektive einer Fortentwicklung von Bildungszielen und Curricula schließlich kommt ein weiterer Gesichtspunkt hinzu. Es ist anzunehmen, dass Vergleichsstudien zu Schülerleistungen faktisch normierende Wirkungen auf das Bildungswesen ausüben, auch wenn dies zunächst nicht intendiert sein sollte. In teils impliziter, teils expliziter Weise hat die internationale und nationale Bildungsdiskussion vielerorts begonnen, sich an den Kompetenzdefinitionen zu orientieren, die durch die jeweiligen Aufgabenserien dieser Studien bereitgestellt werden. Sinnvoll schiene es mir, solche Rückwirkungen von Evaluationsstudien empirisch zu erfassen und bildungstheoretisch angemessen zu bewerten.

### *3.7 Erweiterung der Perspektiven II: Möglichkeitsräume des Bildungswesens – Forschung zum schulischen Bildungswesen anderer Länder und Epochen*

Internationale Vergleichsstudien zu Schülerleistungen sind per definitionem an internationalen und interkulturellen Vergleichsmaßstäben orientiert. Dennoch drängt sich der Eindruck auf, dass selbst komparativ orientierte Forschung dieser Art heute typischerweise zunächst unter dem Blickwinkel national spezifizierter Fragestellungen betrieben und interpretiert wird (wie erfolgreich ist das nationale Bildungswesen im Vergleich mit den Bildungssystemen anderer Länder, wie lassen sich Leistungsrankpositionen der eigenen Schülerpopulationen verbessern etc.). Spezifizierungen dieser Art sind sinnvoll und legitim. Wissenschaftlich und bildungspolitisch schiene es mir aber ebenfalls wesentlich zu sein, den Geltungsanspruch einer schulbezogenen Bildungsforschung über nationale Ansprüche hinaus programmatisch zu erweitern.

Wesentlich könnte insbesondere Forschung zum Aufbau des Bildungswesens in denjenigen Ländern sein, denen es bisher an einer hinreichenden eigenen bildungswissenschaftlichen Infrastruktur mangelt. Ferner könnte es sich als hilfreich erweisen, wenn – in Koordination mit den gegenwärtigen Bemühungen der OECD – von deutscher Seite Maßnahmen zur Entwicklung von Forschungsstrukturen in solchen Ländern initiiert würden (in Gestalt von auslandsorientierten bildungswissenschaftlichen Studiengängen an deutschen Universitäten, Programmen des Wissenschaftleraustauschs etc.).

Neben komparativen Analysen zu gegenwärtigen Realisierungen von Bil-



dungssystemen könnte sich systematische Forschung zu historischen Realisierungen und Entwicklungen des Bildungswesens in unterschiedlichen Ländern als fruchtbringend erweisen. Unter kulturanthropologischen Perspektiven ist solche Forschung Voraussetzung einer Einordnung des Stellenwerts von Bildungssystemen für die kulturelle Evolution des Menschen. Auch unter bildungswissenschaftlicher Perspektive aber dürfte historisch-vergleichende Forschung unverzichtbar sein, wenn der Möglichkeitsraum denkbarer Realisierungen von Bildungssystemen sichtbar werden soll: Für Konzeptionen eines solchen Möglichkeitsraums werden Erkenntnisse zu Realisierungen in vergangenen Epochen und in der Gegenwart unterschiedlicher Kulturen ebenso benötigt wie Einsichten der Interventionsforschung zu zukünftig möglichen Unterrichts- und Systemformen.

Wenn es gelänge, Perspektiven kulturkomparativer, historisch-komparativer und gestaltend-interventiver Analysen zu integrieren, könnten günstigere Voraussetzungen für eine empirisch fundierte Abschätzung möglicher, wünschenswerter und zu verhindernder Weiterentwicklungen des Bildungswesens gewonnen werden. Ähnlich wie beispielsweise im Bereich der Technikfolgen-Abschätzung könnte es im Zuge von Prospektionsstudien dann auch möglich werden, antizipatorische Analysen der Folgen und Nebenwirkungen zukünftiger Entwicklungen des Bildungswesens durchzuführen, um Bildungspolitik und pädagogischer Praxis Entscheidungsgrundlagen für langfristige-zukunftsorientiertes Handeln zu liefern.

#### **4. Desiderata zu Methoden und Organisation der schulbezogenen Bildungsforschung**

Die bisherige schulbezogene Bildungsforschung hat nicht durchweg immer denjenigen Ertrag erbracht, der für kumulativen wissenschaftlichen Fortschritt und damit auch für bildungspolitisches und pädagogisch-praktisches Handeln erforderlich scheint. Mithin ergeben sich auch zu den Methoden und Organisationsformen einer zukünftigen schulbezogenen Bildungsforschung eine Reihe von Desiderata.

##### *4.1 Forschungsmethoden und Leitlinien für gute Forschung*

International scheint heute in manchen Bereichen der Bildungsforschung ein Trend zu dominieren, an Einzelfällen orientierte Forschungsstrategien narrativ-qualitativer Art zu präferieren (insbesondere in den USA). Probleme solcher Strategien liegen nicht nur im zuweilen vertretenen Dominanzanspruch

ihrer Protagonisten. Vielmehr scheinen solche Strategien häufig mit einem Verlust jeglichen expliziten Anspruchs auf Generalisierbarkeit der jeweiligen fallbezogenen Forschungsbefunde einherzugehen (vgl. Eisner 1997; Howe 1998; Mayer 2000). Befunde aber, die über den untersuchten Fall hinaus nicht generalisierbar sind, können weder zu kumulativem Erkenntnisfortschritt beitragen noch lassen sie sich praktisch nutzen. Ausnahmen ergeben sich nur in denjenigen Bereichen, in denen bereits das Forschungsfeld vorwiegend einzelfallartig organisiert ist und insoweit einer primär idiographischen Methodik bedarf (wie z.B. in der historischen Bildungsforschung).

Für die europäische und deutsche Bildungsforschung scheint mir deshalb entscheidend zu sein, amerikanischen Vorbildern dieser Art nicht zu folgen. Vielmehr ist nach wie vor als Grundmerkmal einer als wissenschaftlich zu bezeichnenden, empirischen Bildungsforschung anzusehen, dass sich verallgemeinerbare und insoweit auch nutzbare Schlüsse ergeben.

Im Sinne einer selbstorganisierten Qualitätssicherung der empirischen Bildungsforschung könnten aus dieser Zielstellung *Leitlinien für gute Forschungspraxis* abgeleitet werden, an denen sich die Entwicklung und Evaluation von Forschungsprogrammen orientieren könnte. Solche Leitlinien könnten dafür sorgen, dass unterschiedlichen Strategien der Erkenntnisgewinnung der jeweils heuristisch angemessene Platz eingeräumt wird. Dies böte gleichzeitig auch die Chance, von der Bildungsforschung bisher unzureichend genutzten Strategien einen größeren Stellenwert einzuräumen. Zwei Beispiele sind (a) explorative Strategien zur theorieoffenen Erkundung neuer Gegenstandsbereiche sowie (b) idiographisch-nomothetische Strategien einer Analyse von Unterschieden und Gemeinsamkeiten zwischen Einzelfällen (Individuen, Einzelschulen etc.), die auf nomothetisch generalisierbare Schlussfolgerungen zielen.

Strategien beiderlei Art sind bisher vernachlässigt worden. Insbesondere ist festzustellen, dass traditionelle Forschungsstrategien stichproben- und populationsorientierter Art allein nicht sicherstellen können, dass den jeweiligen Schlussfolgerungen tatsächlich Validität für einzelne Personen bzw. Organisationseinheiten zukommt. Methodisch ist trivial, dass sich aus Stichproben- und Populationsstatistiken Beschreibungsmerkmale von Stichproben bzw. Populationen gewinnen lassen, hingegen keine Schlüsse zu individuellen Entwicklungs- und Funktionsbeziehungen ableitbar sind (Pekrun u.a. im Druck). Tatsächlich aber verhalten sich weite Bereiche von psychologischer und pädagogischer Bildungsforschung nach wie vor so, als seien solche Schlüsse möglich (Beispiel: Schlüsse auf die Leistungswirksamkeit von Schüler selbstkonzepten, die aus korrelativ-interindividuellen Stichprobenzusammenhängen von Selbstkonzept- und Leistungswerten gezogen werden). Notwendig schiene mir hier ein methodischer Paradigmenwechsel hin zu

denjenigen in der rezenten Methodenforschung entwickelten Strategien, die systematische, intraindividuelle Einzelfallanalysen so anlegen und miteinander verknüpfen, dass nomothetische Schlussfolgerungen möglich werden, denen tatsächlich *generelle* wie *individuelle* Validität zuzusprechen ist (Schmitz 2000).

#### 4.2 Organisation von Forschung: Interdisziplinäre Kooperation und Stärkung der Fachdidaktiken

Die meisten Gegenstandsbereiche der schulbezogenen Bildungsforschung bedürfen eines interdisziplinär organisierten Zugriffs. Dies gilt beispielsweise für den Wissenserwerb in einzelnen Schulfächern, für dessen Analyse eine Zusammenarbeit von Wissenspsychologie, empirischer Pädagogik und Fachdidaktik notwendig scheint. Eine solche Zusammenarbeit ist bisher nicht hinreichend häufig realisiert worden, hat sich aber z.B. in den Schülerleistungsstudien der letzten Jahren (z.B. PISA) als außerordentlich fruchtbar erwiesen. Ein zweites Beispiel ist die komparative Analyse von Bildungssystemen, für die eine engere Zusammenarbeit von kultur- und epochenvergleichender Pädagogik, Psychologie, Bildungssoziologie und Bildungsökonomie hilfreich wäre.

Voraussetzung einer solchen Zusammenarbeit wäre, dass alle beteiligten Disziplinen über die traditionelle Verwendung analytischer, normativer und konstruktiver Strategien hinaus gleichzeitig auch konsequent empirisch orientiert vorgehen. In einer Reihe von fachdidaktischen Disziplinen scheint es bisher auch in Deutschland teilweise an einer solchen Orientierung zu mangeln. Im Sinne der Gesamtziele schulbezogener Bildungsforschung dürfte es deshalb wesentlich sein, entsprechende Entwicklungen in den betreffenden Disziplinen zu verstärken bzw. in die Wege zu leiten. Die mancherorts geplante oder bereits realisierte Reduktion wissenschaftlicher Personalkapazität der Fachdidaktiken hingegen könnte sich unter Gesichtspunkten eines gezielten Ausbaus der deutschen Bildungsforschung als hinderlich erweisen. Vielmehr bedürfte es m.E. einer auch quantitativen Stärkung der betreffenden Disziplinen an Universitäten wie außeruniversitären Forschungseinrichtungen.

Für den Ausbau einer empirischen, interdisziplinär orientierten Bildungsforschung in Deutschland stehen geeignete Instrumente der Forschungsförderung und forschungsbezogenen Nachwuchsförderung im Grundsatz zur Verfügung (Sonderforschungsbereiche, Forschungszentren, Schwerpunktprogramme der DFG, interdisziplinäre Forschungsverbünde der Länder etc.). Es käme darauf an, diese Instrumente nun auch seitens der Lehr-Lern- und Bildungsforschung in verstärktem Maße zu nutzen.

**Literatur**

- Anderson, J.R./Redder, L.M./Simon, H.A. (1997): Situative versus cognitive perspectives: Form versus substance. In: *Educational Researcher* 26, H. 1, S. 18-21.
- Baumert, J./Bos, W./Lehmann, R. (Hrsg.) (2000): TIMSS/III: Dritte internationale Mathematik- und Naturwissenschaftsstudie: Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn, 2 Bde. Opladen. Leske & Budrich.
- Baumert, J./Lehmann, R. (1997): TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Opladen: Leske & Budrich.
- Benner, D. (2002): Die Struktur der Allgemeinbildung im Kerncurriculum moderner Bildungssysteme. Ein Vorschlag zur bildungstheoretischen Rahmung von PISA. In: *Zeitschrift für Pädagogik* 48, S. 68-90.
- Ceci, S.J. (1991): How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. In: *Developmental Psychology* 27, S. 703-722.
- Cobb, P./Bowers, J. (1999): Cognitive and situated learning perspectives in theory and practice. *Educational Researcher* 28, H. 2, S. 4-15.
- Eisner, E.W. (1997): The promise and perils of alternative forms of data representation. In: *Educational Researcher* 26, H. 6, S. 4-10.
- Fend, H./Stöckli, G. (1997): Der Einfluss des Bildungssystems auf die Humanentwicklung: Entwicklungspsychologie der Schulzeit. In: Weinert, F.E. (Hrsg.): *Psychologie des Unterrichts und der Schule (Enzyklopädie der Psychologie, Serie Pädagogische Psychologie, Bd. 3)*. Göttingen: Hogrefe, S. 1-35.
- Gerstenmaier, J./Mandl, H. (1995): Wissenserwerb unter konstruktivistischer Perspektive. In: *Zeitschrift für Pädagogik* 41, S. 867-888.
- Gruber, H./Ziegler, A. (Hrsg.) (1996): *Expertiseforschung*. Opladen: Leske & Budrich.
- Helmke, A./Weinert, F.E. (1997): Bedingungsfaktoren schulischer Leistungen. In: Weinert, F.E. (Hrsg.): *Psychologie des Unterrichts und der Schule (Enzyklopädie der Psychologie, Serie Pädagogische Psychologie, Bd. 3)*. Göttingen: Hogrefe, S. 71-176.
- Howe, K.R. (1998): The interpretive turn and the new debate in education. In: *Educational Researcher* 27, H. 8, S. 13-22.
- Klieme, E./Artelt, C./Stanat, P. (2001): Fächerübergreifende Kompetenzen: Konzepte und Indikatoren. In: Weinert, F.E. (Hrsg.): *Leistungsmessungen in Schulen*. Weinheim: Beltz, S. 203-218.
- Kline, R.B. (1998): *Principles and practice of structural equations modeling*. New York: Guilford Press.
- Klix, F./Spada, H. (Hrsg.) (1998): *Wissen (Enzyklopädie der Psychologie, Serie Kognition, Bd. 6)*. Göttingen: Hogrefe.
- Krapp, A. (1993): Die Psychologie der Lernmotivation. In: *Zeitschrift für Pädagogik* 39, S. 187-206.
- Loehlin, J.C./McCrae, R.R./Costa, P.T./John, O.P. (1998): Heritabilities of common and measure-specific components of the Big Five personality factors. In: *Journal of Research in Personality* 32, S. 431-453.
- Lynn, R. (1997): Intelligence in Taiwan. In: *Personality and Individual Differences* 22, S. 585-586.
- Mandl, H./Reinmann-Rothmeier, G./Graesel, C. (1998): Gutachten zum BLK-Programm: Systematische Einbeziehung von Medien, Informations- und Kommunikationstechnologien in Lehr-Lern-Prozesse. München: Fakultät für Psychologie und Pädagogik.

- Mayer, R.E. (2000): What is the place of science in educational research? In: *Educational Researcher* 29, H. 6, S. 38-39.
- Pekrun, R. (1997). Kooperation zwischen Schule und Elternhaus. In: Vaskovics, L.A./Lipinski, H. (Hrsg.): *Familiale Lebenswelten und Bildungsarbeit*. Opladen: Leske & Budrich, S. 51-79.
- Pekrun, R. (im Druck): Familie, Schule und Entwicklung. In: Walper, S./Pekrun, R. (Hrsg.): *Familie und Entwicklung*. Göttingen: Hogrefe.
- Pekrun, R./Fend, H. (Hrsg.) (1991): *Schule und Persönlichkeitsentwicklung. Ein Resümee der Längsschnittforschung*. Stuttgart: Enke.
- Pekrun, R./Götz, T./Titz, W./Hofmann, H. (im Druck): Mythen in der Schulleistungs- und Jugendforschung: Möglichkeiten und Grenzen populationsorientierter Untersuchungen. In: Merckens, H./Zinnecker, J. (Hrsg.): *Jahrbuch Jugendforschung* 2. Opladen: Leske & Budrich.
- Pekrun, R./Hofmann, H. (1999): Lern- und Leistungsemotionen: Erste Befunde eines Forschungsprogramms. In: Jerusalem, M./Pekrun, R. (Hrsg.): *Emotion, Motivation und Leistung*. Göttingen: Hogrefe, S. 247-267.
- Perry, R./Smart, J.C. (Hrsg.) (1997): *Effective teaching in higher education: Research and practice*. New York: Agathon Press.
- Reusser, K. (1998): Denkstrukturen und Wissenserwerb in der Ontogenese. In: Klix, F./Spada, H. (Hrsg.): *Wissen (Enzyklopädie der Psychologie, Serie Kognition, Bd. 6)*. Göttingen: Hogrefe, S. 115-166.
- Rolff, H.-G. (1996). Verwaltungsreform im Schulbereich: Schlanker und pädagogischer zugleich? Ein Fallbeispiel aus Bremen. In: *Organisationsentwicklung* 15, S. 30-40.
- Schibeci, R.A./Grundy, S. (1987): Local theories. In: *Journal of Educational Research* 81, S. 91-96.
- Schiefele, U. (im Druck): The role of interest in motivation and learning. In: Messick, S./Collis, J.M. (Hrsg.): *Intelligence and personality. Bridging the gap in theory and measurement*. Hillsdale: Lawrence Erlbaum.
- Schmitz, B. (2000): Auf der Suche nach dem verlorenen Individuum: Vier Theoreme zur Aggregation von Prozessen. In: *Psychologische Rundschau* 51, S. 83-92.
- Strittmatter, P./Bedersdorfer, H.W. (1991): Pädagogische Interventionsforschung: Abbau von Angst in schulischen Leistungssituationen. In: Pekrun, R./Fend, H. (Hrsg.): *Schule und Persönlichkeitsentwicklung. Ein Resümee der Längsschnittforschung*. Stuttgart: Enke, S. 297-323.
- Tergan, S.O. (1998): Checklists for the evaluation of educational software: Critical review and prospects. In: *Innovations in Education and Training International* 35, S. 9-20.
- Terhart, E. (Hrsg.) (2000): *Perspektiven der Lehrerbildung in Deutschland. Abschlussbericht der von der Kultusministerkonferenz eingesetzten Kommission*. Weinheim: Beltz.
- Terhart, E. (2002): Wie können die Ergebnisse von vergleichenden Leistungsstudien systematisch zur Qualitätsverbesserung in Schulen genutzt werden? In: *Zeitschrift für Pädagogik* 48, S. 91-110.
- Thomas, A. (1996): *Psychologie interkulturellen Handelns*. Göttingen: Hogrefe.
- Weinert, F.E. (Hrsg.) (2001): *Leistungsmessungen in Schulen*. Weinheim: Beltz.
- Weinert, F.E. (1998): Vermittlung von Schlüsselqualifikationen. In: Matalik, S./Schade, D. (Hrsg.): *Entwicklungen in Aus- und Weiterbildung*. Baden-Baden, Akademie für Technikfolgenabschätzung, S. 23-43.
- Wottawa, H./Thierau, H. (1998): *Lehrbuch Evaluation*. Bern: Huber.

**Abstract:** *International and national studies on the evaluation of educational achievement provide data on distributions, structures, and correlates of students' competences and achievement. This implies that they may serve functions of a descriptive monitoring of school systems. In contrast to many expectations by politicians and the public, however, their cross-sectional designs do not allow any conclusions on the antecedents and consequences of students' achievement, or on educational interventions. Implications of the benefits and limitations of such studies are discussed. One implication is that more programs of developmental, causal, and intervention-oriented research on students' education should be implemented in Germany. In addition, consequences for the organization and methodology of such programs are discussed.*

*Anschrift des Autors:*

Prof. Dr. Reinhard Pekrun, Institut für Pädagogische Psychologie, Universität München,  
Leopoldstrasse 13, 80802 München.